

Міністерство освіти і науки України
Національний аерокосмічний університет
«Харківський авіаційний інститут»

Кафедра математичного моделювання та штучного інтелекту (№ 304)

ЗАТВЕРДЖУЮ

Гарант освітньої програми



(підпис)

Юрій СКОБ

(ім'я та ПРІЗВИЩЕ)

« 21 » 08 2025 р.

**СИЛАБУС ОBOB'ЯЗKОВОЇ
НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

Інтелектуальний аналіз даних

(назва навчальної дисципліни)

Галузь знань: F «Інформаційні технології»

(шифр і найменування галузі знань)

Спеціальність: F3 «Комп'ютерні науки»

(код і найменування спеціальності)

Освітня програма: «Інтелектуальні системи та технології»

(найменування освітньої програми)

Рівень вищої освіти: *другий (магістерський)*

Силабус уведено в дію з 01.09.2025

Харків – 2025 р.

Розробник: Базілевич К.О., доц. каф. 304, к. т. н., доц.
(прізвище та ініціали, посада, науковий ступінь і вчене звання)


(підпис)

Силабус навчальної дисципліни розглянуто на засіданні кафедри № 304
математичного моделювання та штучного інтелекту
(назва кафедри)

Протокол № 1 від « 21 » серпня 2025 р.

В. о. завідувача кафедри к.ф.-м.н., доцент
(науковий ступінь і вчене звання)


(підпис)

Олексій КАРТАШОВ
(ім'я та ПРІЗВИЩЕ)

Погоджено з представником здобувачів освіти:

Здобувач другого (магістерського)

рівня вищої освіти за ОП

«Інтелектуальні системи та технології»


(підпис)

Анастасія КОЗАКОВА
(ім'я та ПРІЗВИЩЕ)

1. Загальна інформація про викладача



ПІБ: *Базілевич Ксенія Олексіївна*

Посада: *доцент каф. 304*

Науковий ступінь: *канд. техн. наук*

Вчене звання: *доцент*

Перелік дисциплін, які викладає:

Інтелектуальний аналіз даних; Проєктування програмного забезпечення; Сучасні методи проєктування програмного забезпечення

Напрями наукових досліджень:

Інтелектуальний аналіз даних

Контактна інформація:

k.bazilevych@khai.edu

2. Опис навчальної дисципліни

Форма здобуття освіти	Денна
Семестр	2
Мова викладання	Українська
Тип дисципліни	Обов'язкова
Обсяг дисципліни: кредити ЄКТС/ кількість годин	5 кредитів ЄКТС / 150 годин (64 аудиторних, з яких: лекції – 32, лабораторні – 32; СРЗ – 86);
Види навчальної діяльності	Лекції та лабораторні заняття, самостійна робота
Види контролю	Поточний контроль, модульний контроль, семестровий контроль – іспит
Пререквізити	«Теорія алгоритмів і математична логіка», «Методи обчислень», «Об'єктно-орієнтоване програмування (мова C#)»

3. Мета та завдання навчальної дисципліни, переліки компетентностей та очікуваних результатів навчання

Мета – поглиблення знань та ефективне використання сучасних методів інтелектуального аналізу даних та обробки інформації з використанням програмного забезпечення.

Завдання – оволодіння сучасною методологією та особливостями застосування інтелектуальної обробки даних; засвоєння майбутніми фахівцями теоретичних основ інформаційних систем, орієнтованих на застосування стандартів інтелектуального аналізу даних; отримання практичних навичок використання і адаптації деяких найбільш відомих систем та бібліотек для реалізації інтелектуального аналізу даних.

Компетентності, які набуваються:

Інтегральна компетентність:

Здатність розв'язувати задачі дослідницького та/або інноваційного характеру у сфері комп'ютерних наук.

Загальні компетентності (ЗК)

Після закінчення цієї програми здобувач освіти буде здатен:

- до абстрактного мислення, аналізу та синтезу (ЗК1);
- застосовувати знання у практичних ситуаціях (ЗК2);
- вчитися й оволодівати сучасними знаннями (ЗК 5).

Спеціальні компетентності (СК)

Після закінчення цієї програми здобувач освіти буде здатен:

- використовувати математичні методи для аналізу формалізованих моделей предметної області (СК 3);
- збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проєктних рішень (СК 4);
- застосовувати існуючі і розробляти нові алгоритми розв'язування задач у галузі комп'ютерних наук (СК6).

Програмні результати навчання (ПРН):

- Розробляти математичні моделі та методи аналізу даних (включно з великим) (ПРН 8);
- Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими) (ПРН 9).

4. Зміст навчальної дисципліни

МОДУЛЬ 1

Змістовий модуль 1. Основні задачі інтелектуального аналізу даних

Тема 1. Визначення, основні задачі інтелектуального аналізу даних (ІАД). Машинне навчання.

Стисла анотація: Інтелектуальний аналіз даних. Визначення Data Mining. Методи і задачі Data Mining. Зв'язки з іншими областями знань. Класифікація стадій Data Mining. Поняття Data Science. Аналіз даних та інтелектуальний аналіз даних. Приклади рішень, продуктів які створені за допомогою Data Science. Основні області застосування Data Science. Використання Machine learning.

Теми лекцій: «Визначення, основні задачі інтелектуального аналізу даних (ІАД). Машинне навчання».

Теми лабораторних занять: «Моделювання як науковий метод».

Самостійна робота здобувача освіти: опрацювання матеріалу лекцій, виконання індивідуальних завдань відповідно варіанту, підготовка до захисту лабораторних робіт.

Тема 2. Препроцесінг інформації

Стисла анотація: Цикл одержання попередньої обробки аналізу даних, інтерпретації результатів та їхнього використання. Ентропія і кількість інформації. Нормалізація і стандартизація вихідних значень. Алгоритм «вибілювання» входів. Виключення неінформативних факторів. Аналітико-евристичні алгоритми визначення вагомих інформативних ознак.

Теми лекцій і лабораторних занять: «Оцінка інформативності факторів».

Самостійна робота здобувача освіти: опрацювання матеріалу лекцій, виконання індивідуальних завдань відповідно варіанту, підготовка до захисту лабораторних робіт.

Модульний контроль 1

Змістовий модуль 2. Використання методів Data Mining та машинного навчання

Тема 3. Задача класифікації

Стисла анотація: Завдання Data Mining: класифікація задач Data Mining, завдання класифікації і регресії, завдання пошуку асоціативних правил. Постановка задачі класифікації. Представлення результатів: правила класифікації, дерева рішень, математичні функції. Методи опорних векторів, «найближчого сусіда», алгоритм побудови 1-правил, метод Naive Bayes. Метод Random Forest. Аналіз багатомірних угруповань. Класифікація об'єктів у випадку невідомих розподілень даних. Методи оцінювання помилок класифікації. Оцінка параметрів логістичної регресії на основі методу оцінки шансів і ймовірностей. Оцінка ймовірності за

допомогою байєсовського класифікатора (Naive Bayes). Задача прогнозування.

Теми лекцій і лабораторних занять: «Методи класифікації».

Самостійна робота здобувача освіти: опрацювання матеріалу лекцій, виконання індивідуальних завдань відповідно варіанту, підготовка до захисту лабораторних робіт.

Тема 4. Задача кластеризації

Стисла анотація: Базові алгоритми кластеризації. Адаптивні методи кластеризації. Постановка завдання кластеризації: формальна постановка задачі, заходи близькості, засновані на відстанях, які використовуються в алгоритмах кластеризації. Представлення результатів кластеризації. Види кластерів. Базові алгоритми кластеризації: класифікація алгоритмів, ієрархічні алгоритми, неієрархічні алгоритми. Адаптивні методи кластеризації: вибір найкращого рішення і якості кластеризації, використання формальних критеріїв якості в адаптивній кластеризації, приклад адаптивної кластеризації. Метод ліктя. Оцінка якості кластеризації.

Теми лекцій і лабораторних занять: «Методи кластеризації».

Самостійна робота здобувача освіти: опрацювання матеріалу лекцій, виконання індивідуальних завдань відповідно варіанту, підготовка до захисту лабораторних робіт.

Тема 5. Математичні методи аналізу часових рядів

Авторегресійні моделі, моделі ковзного середнього та їх комбінації (модель Бокса-Дженкінса підбору ARIMA моделі для часового ряду, SARIMA). Використання машинного навчання. Нелінійна адитивна модель у модулі Facebook Prophet.

Теми лекцій і лабораторних занять: «Методи аналізу часових рядів».

Самостійна робота здобувача освіти: опрацювання матеріалу лекцій, виконання індивідуальних завдань відповідно варіанту, підготовка до захисту лабораторних робіт.

Модульний контроль 2

5. Індивідуальні завдання

Робота розрахункова (РР) «Проведення власного дослідження із застосуванням методів інтелектуального аналізу даних»

1. Обрати статистичні дані з будь-якої предметної області для аналізу. Приклади додаються в файлі з завданням.

2. Обрати самостійно мету дослідження: сформулювати постановку задачі та хід дослідження, з урахуванням того, яка задача буде вирішуватись.

3. Провести розрахунки за допомогою трьох методів інтелектуального аналізу даних.
4. Розробити алгоритмічну модель та блок-схему кожного методу.
5. Розробити програмний додаток, що вирішує описані задачі та підтримує завантаження статичних даних для розрахунків.
6. Порівняти результати розрахунків
6. Оформити звіт по виконаних завданнях, зробити висновки.

6. Методи навчання

1. Пояснювально-ілюстративний (інформаційно-рецептивний) метод: та метод проблемного виконання (лекція).
2. Репродуктивний (лабораторні роботи).
3. Частково-пошуковий (евристичний) та дослідницький: (самостійна робота).
4. Дисципліна «Інтелектуальний аналіз даних» передбачає лекційні (в т. ч. з використанням мультимедійного обладнання) заняття і виконання лабораторних робіт під керівництвом викладача та самостійну роботу студента за підручниками і матеріалами, що забезпечує закріплення теоретичних знань, сприяє набуттю практичних навичок і розвитку самостійного наукового мислення. Передбачено регулярні індивідуальні консультації.

7. Методи контролю

Визначення рівня засвоєння студентом навчального матеріалу дисципліни здійснюється шляхом проведення поточних (захист лабораторних робіт) і підсумкових контролів (захист змістовного модуля, іспит).

Поточний контроль має на меті систематично відстежувати, наскільки студент розуміє та засвоює матеріал, вміє виконувати лабораторні роботи, самостійно вирішувати завдання та писати відповідний код, готувати звіти та представляти результати дослідження в усній або письмовій формі.

Підсумковий контроль покликаний оцінити глибину знань студента з усього курсу, його розуміння зв'язків між різними темами, здатність застосовувати отримані знання на практиці та самостійно вирішувати задачі з інтелектуального аналізу даних, які виникають в рамках дисципліни.

8. Критерії оцінювання та розподіл балів, які отримують здобувачі освіти

Таблиця 8.1 – Розподіл балів, які отримують здобувачі освіти

Складові навчальної роботи	Бали за одне заняття (завдання)	Кількість занять (завдань)	Сумарна кількість балів
Змістовний модуль 1			
Виконання і захист лабораторних робіт	0...15	2	0...30
Модульний контроль	0...5	2	0...10
Змістовний модуль 2			
Виконання і захист лабораторних	0...15	3	0...45
Модульний контроль	0...5	2	0...10
Виконання і захист РК	0...5	1	0...5
Усього за семестр			0...100

Семестровий контроль (*іспит*) проводиться у разі відмови здобувача освіти від балів підсумкового контролю й за наявності допуску до *іспиту*. Під час складання семестрового *іспиту* здобувач освіти має можливість отримати максимум 100 балів. Білет для *іспиту* складається з одного теоретичного і двох практичних питань. У першому запитанні студент має продемонструвати теоретичні знання. У другому запитанні – показати навички складання і виконання комп’ютерного додатку для вирішення задачі аналізу даних за заданою темою. У третьому пункті – продемонструвати вміння проводити перевірку отриманих результатів за допомогою стандартних функцій та бібліотек. Максимальна кількість балів за кожне запитання та їх суми показані в табл. 8.2.

Таблиця 8.2 – Шкали оцінювання відповідей на запитання білету

Складові білету	Складові оцінки	Бали за одне питання	Сумарна кількість балів
Пункт 1. Теоретичне запитання	Надано визначення поняття	0...10	0...20
	Наведено власний приклад використання	0...10	
Пункт 2. Практичне запитання	Складено алгоритм розрахунку (блок-схема) задачі	0...15	0...50
	Створено проект додатку для вирішення задачі	0...20	
	Отримано правильні результати розрахунків	0...15	
Пункт 3. Практичне запитання	Проведено перевірку отриманих результатів за допомогою стандартних функцій та бібліотек	0...30	0...30
Підсумкова оцінка за іспит			0...100

Таблиця 8.3 – Шкали оцінювання: бальна і традиційна

Сума балів	Оцінка за традиційною шкалою	
	Іспит, диференційний залік	Залік
90 – 100	Відмінно	Зараховано
75 – 89	Добре	
60 – 74	Задовільно	
0 – 59	Незадовільно	Не зараховано

Критерії оцінювання роботи здобувача освіти протягом семестру

Задовільно (60-74): має базові знання та навички, необхідні для досягнення програмних результатів навчання з інтелектуального аналізу даних; виконав та успішно здав лабораторні роботи, що охоплюють основні задачі Data Mining; написав модульні роботи, що демонструють розуміння ключових концепцій та технік аналізу даних; вміє застосовувати основні інструменти та бібліотеки для попередньої обробки даних, знає застосування базових алгоритмів класифікації та кластеризації, а також вміє проводити їхню оцінку.

Добре (75-89): має знання та навички, достатні для досягнення програмних результатів навчання з інтелектуального аналізу даних; виконав та успішно здав лабораторні роботи, включаючи реалізацію та порівняння різних алгоритмів класифікації та кластеризації; написав модульні роботи, що демонструють глибоке розуміння концепцій та технік аналізу даних. Додатково до вимог для оцінки «Задовільно»: вміння обґрунтовано вибирати відповідні алгоритми класифікації та кластеризації для розв'язання конкретних задач аналізу даних; вміння проводити експерименти з різними параметрами алгоритмів та інтерпретувати результати; вміння критично оцінювати результати аналізу даних та пропонувати шляхи для покращення моделей; вміння візуалізувати результати аналізу даних, використовуючи відповідні графіки та діаграми для їхньої інтерпретації..

Відмінно (90-100): володіє глибокими знаннями, вміннями та навичками, що дозволяють самостійно та обґрунтовано розв'язувати складні задачі аналізу даних; досконале виконання та захист всіх лабораторних робіт; успішно написані модульні роботи, що демонструють здатність до самостійного дослідження та аналізу проблем у сфері аналізу даних; вміння самостійно використовувати широкий спектр інструментів та бібліотек для повного циклу аналізу даних: від збору та попередньої обробки до побудови та оцінки складних моделей (включаючи часові ряди); вміння розробляти та реалізовувати власні алгоритми або модифікації існуючих алгоритмів для вирішення специфічних задач; вміння ефективно використовувати методи зменшення розмірності даних, виявлення аномалій та інші передові техніки; вміння представляти результати аналізу даних у вигляді чітких та інформативних звітів, що містять обґрунтовані висновки та рекомендації.

9. Політика навчального курсу

Відвідування занять. З огляду на інтерактивний формат курсу, відвідування практичних занять є обов'язковим. Студенти, які не можуть регулярно відвідувати практичні заняття, на яких виконуються лабораторні роботи, повинні протягом тижня узгодити з викладачем індивідуальний графік відпрацювання пропущених занять. Пропущені заняття відпрацьовуються на найближчій консультації протягом тижня після пропуску у формі співбесіди за питаннями з плану заняття. В окремих випадках дозволяється відпрацювання у письмовій формі шляхом виконання індивідуального письмового завдання.

Дотримання вимог академічної доброчесності здобувачами освіти під час вивчення навчальної дисципліни. Під час вивчення навчальної дисципліни здобувачі освіти мають дотримуватися загальноприйнятих морально-етичних норм і правил поведінки, вимог академічної доброчесності, передбачених Положенням про академічну доброчесність Національного аерокосмічного університету «Харківський авіаційний інститут» (<https://khai.edu/assets/files/polozhennya/polozhennya-pro-akademichnu-dobrochesnist.pdf>). Очікується, що роботи здобувачів освіти будуть їх оригінальними дослідженнями або міркуваннями. Відсутність посилань на використані джерела, фабрикавання джерел, списування, втручання в роботу інших здобувачів освіти становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в письмовій роботі здобувача освіти є підставою для її незарахування викладачем незалежно від масштабів плагіату чи обману.

Вирішення конфліктів. Порядок і процедури врегулювання конфліктів, пов'язаних із корупційними діями, зіткненням інтересів, різними формами дискримінації, сексуальними домаганнями, міжособистісними стосунками та іншими ситуаціями, що можуть виникнути під час навчання, а також правила етичної поведінки регламентуються Кодексом етичної поведінки в Національному аерокосмічному університеті «Харківський авіаційний інститут» (<https://khai.edu/ua/university/normativna-baza/ustanovchi-dokumenti/kodeks-etichnoi-povedinki/>).

10. Методичне забезпечення

Увесь науково методичний комплект з дисципліни розміщено на освітньому порталі Mentor: <https://mentor.khai.edu/course/view.php?id=7667>.

11. Рекомендована література

Базова

1. Інтелектуальний аналіз даних : навчальний посібник / А. О. Олійник, С. О. Субботін, О. О. Олійник. – Запоріжжя : ЗНТУ, 2012. – 278 с.
2. Ian H. Witten. Data Mining: Practical Machine Learning Tools and Techniques / Ian H. Witten, Eibe Frank and Mark A. Hall — [3rd Edition] — Morgan Kaufmann, 2011. — Р. 664. — ISBN 9780123748560.
3. Гороховатський В.О., Творошенко І.С. Методи інтелектуального аналізу та оброблення даних: навч. посібник. – Харків: ХНУРЕ, 2021. – 92 с.
4. Encyclopedia of artificial intelligence / Eds.: J. R. Dopico, J. D. de la Calle, A. P. Sierra. –New York : Information Science Reference, 2009. – Vol. 1–3. – 1677 p.
5. Ситник В.Ф. Засоби дейтамайнінгу для аналізу бізнесових рішень. Науково-практичний журнал "Науково-технічна інформація", №3, 2002. - с. 60-64.
6. Ситник В.Ф. Системи підтримки прийняття рішень: Навч.посібник / В. Ф. Ситник - К.:КНЕУ, 2004. – 614 с.
7. Yunida H. Saving of Maternal and Infant Lives with Sustainable Midwifery Services. Int J Community Based Nurs Midwifery (2022) 313-314. doi: 10.30476/IJCBNM.2022.95877.2092.
8. Goodfellow I. Deep Learning (2018) 652 с.
9. Закон України «Про вищу освіту» від 01.07.2014 р. № 1556-VII // *Відомості Верховної Ради України*. – 2014. – № 37–38. – Ст. 2004. Режим доступу: <https://zakon.rada.gov.ua/laws/show/1556-18>

Допоміжна

1. Єремєєв В.С. Теорія ймовірностей та математична статистика / В.С. Єремєєв, Д.О. Сосновських, О.В. Тітова. - Навчальний посібник: Мелітополь: ТОВ «Видавничий будинок ММД», 2009. – 188 с.
2. Гойко О.В. Аналіз сучасного програмного забезпечення для статистичного оброблення й аналізу біомедичних досліджень. місце видання невідоме : Медична інформатика та інженерія., 2012.
3. Kai-Bo D., S. Sathiya K. «Which Is the Best Multiclass SVM Method? An Empirical Study» – Multiple Classifier Systems, LNCS. pp.278–285.

12. Інформаційні ресурси

1. A. Uberoi: Introduction to Dimensionality Reduction, 2023. URL: <https://www.geeksforgeeks.org/dimensionality-reduction/>
2. Foster P, Tom F. Data Science and its Relationship to Big Data and Data-Driven Decision Making (2013). doi: 10.1089/big.2013.1508

3. James N, Hsien H, Matthew B. Machine learning: applications of artificial intelligence to imaging and diagnosis (2019) 111–118. doi: 10.1007/s12551-018-0449-9
4. Chao Deng, M. Zu Guo. A new co-training-style random forest for computer aided diagnosis. URL: <https://kd.nsfc.gov.cn/paper/Download/1000001632658.pdf>.
5. Yanli L, Yourong W. New Machine Learning Algorithm: Random Forest. (2012) 246-252. doi:10.1007/978-3-642-34062-8_32.
6. Sunil R. *Learn How to Use Support Vector Machines (SVM) for Data Science* (2023). URL: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>.
7. Shamsolmoali, P., Kumar Jain, D., Zareapoor, M., Yang, J. and Afshar Alam, M. High-Dimensional Multimedia Classification Using Deep CNN and Extended Residual Units. *Multimedia Tools and Applications*, (2019), 23867-23882 doi: 10.1007/s11042-018-6146-7.
8. Sruthi R. Understand Random Forest Algorithms With Examples (2023) URL: <https://www.simplilearn.com/tutorials/machine-learning-tutorial/random-ke>
9. Random Forest Algorithm in Machine Learning [Електронний ресурс]. – URL: <https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/>
10. All You Need to Know about Gradient Boosting Algorithm [Електронний ресурс]. – URL: <https://medium.com/data-science/all-you-need-to-know-about-gradient-boosting-algorithm-part-1-regression-2520a34a502>
11. What is LSTM - Long Short Term Memory? [Електронний ресурс]. – URL: <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/>
12. Time series [Електронний ресурс]. – URL: <https://www.sciencedirect.com/topics/economics-econometrics-and-finance/time-series>
13. Time series patterns [Електронний ресурс]. – URL: <https://otexts.com/fpp2/tspatterns.html>
14. Що таке епідемічний поріг, і як його розраховують [Електронний ресурс]. – URL: <https://moz.gov.ua/uk/epidemii-gripu-v-kraini-nemae-scho-treba-znati-pro-perevischennja-epidemichnogo-porogu>
15. ARIMA for Time Series Forecasting [Електронний ресурс]. – URL: <https://www.datacamp.com/tutorial/arima>
16. Model Selection for ARIMA [Електронний ресурс]. – URL: <https://www.geeksforgeeks.org/model-selection-for-arima/>
17. Box-Jenkins Model: Definition, Uses, Timeframes, and Forecasting [Електронний ресурс]. – URL: <https://www.investopedia.com/terms/b/box-jenkins-model.asp>
18. How to Create an ARIMA Model for Time Series Forecasting in Python [Електронний ресурс]. – URL: <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>
19. Tests for Stationarity in Time Series – Dickey Fuller Test & Augmented Dickey Fuller(ADF) Test [Електронний ресурс]. – URL:

<https://medium.com/@ritusantra/tests-for-stationarity-in-time-series-dickey-fuller-test-augmented-dickey-fuller-adf-test-d2e92e214360>

20. Stationarity in Time Series [Электронный ресурс]. – URL: <http://medium.com/@ritusantra/stationarity-in-time-series-887eb42f62a9>

21. What is Time Series and Components of Time Series [Электронный ресурс]. – URL: <https://medium.com/@ritusantra/what-is-time-series-and-components-of-time-series-c80b69ad5cb9>

22. What is a Time Series Unit Root? [Электронный ресурс]. – URL: <https://medium.com/data-science/what-is-a-time-series-unit-root-dba24fa099f4>

23. A Comprehensive Overview of Seasonal ARIMA (SARIMA) Models [Электронный ресурс]. – URL: <https://www.econometricstutor.co.uk/arma-models-seasonal-arma-sarima-models>

24. Introduction to SARIMA Model [Электронный ресурс]. – URL: <https://medium.com/@ritusantra/introduction-to-sarima-model-cbb885ceabe8>

25. Stationarity and differencing [Электронный ресурс]. URL: <https://otexts.com/fpp2/stationarity.html>

26. Facebook Prophet. [Электронный ресурс]. URL: https://facebook.github.io/prophet/docs/quick_start.html#r-api